

Designing Incentive-Compatible Mechanisms for International Climate Cooperation: From the Kyoto Protocol to the Paris Agreement

Zhenglan Tian

*School of Business, Hong Kong Baptist University, Hong Kong, China
25276379@life.hkbu.edu.hk*

Abstract. Climate change has left countries generally facing the dilemma of "hitchhiking" prisoners in emission-reduction cooperation. Taking the Kyoto Protocol and the Paris Agreement as cases, this article compares the incentive effect of the "top-down" compulsory model and the "bottom-up" independent contribution model, and finds three core problems: first, hitchhiking incentives always exist, and countries tend to set too low emission reduction targets; second, the absence of a credible enforcement system renders the cost of non-compliance virtually zero; third, short-term economic benefits continue to overwhelm the long-term climate target. Accordingly, the paper proposes three policy reforms: First, to discourage free-riding, the paper advocates establishing a climate club with a minimum carbon price and border carbon adjustments, supported by graduated sanctions. Second, to address enforcement gaps, it proposes domestic legislation of NDCs, automatic trigger mechanisms, and independent third-party verification. Third, to counter short-termism, it recommends rebalancing costs and benefits through intergenerational transfers, green bonds, and multilateral development bank financing linked to employment and energy security. The social value of this article lies in providing an operational institutional reference for international climate negotiations, and its commercial value lies in reducing the uncertainty around corporate policies and promoting the development of low-carbon industries.

Keywords: International Climate Governance, Incentive-Compatible Mechanisms, Climate Clubs, Enforcement Credibility, Carbon Pricing

1. Introduction

1.1. Research context and problem statement

Climate change represents a global collective action problem of unprecedented scale. According to the IPCC's sixth assessment report, the global surface temperature has risen by about 1.1°C from the pre-industrial level, and extreme weather events are becoming more frequent and intense [1]. As a global public good, climate cooperation is inherently non-excludable and non-rivalrous, giving rise to a classic prisoner's dilemma. Rational, self-interested states tend to under-emit or free-ride,

enjoying the climate stability generated by others' mitigation efforts while avoiding domestic abatement costs.

However, when all countries follow this logic, the collective result is second-best. The "land-in-the-land tragedy" of climate change has become an ongoing real crisis with huge economic and social costs. Therefore, it is of great practical significance to study how to change countries' income structures and guide cooperative behaviour by designing effective incentive mechanisms, thereby resolving the prisoner's dilemma in international climate cooperation. In terms of social value, an effective incentive mechanism serves as a reference for the international multilateral cooperation mechanism. It is a necessary prerequisite for avoiding climate disasters and maintaining the global living environment. In terms of commercial value, it will also promote the vigorous development of green technology, carbon markets and low-carbon industries.

1.2. Theoretical debates and institutional evolution

Scholars have extensively discussed international climate cooperation from the perspective of game theory. Nordhaus put forward the "climate club" theory, arguing that without external sanctions against non-participants, a stable climate cooperation alliance cannot form, and a non-cooperative equilibrium will dominate international climate policy. He proposed changing the game's income structure by establishing a climate club, setting a unified lower limit on carbon prices by member states, and imposing punitive tariffs on non-member states [2]. By contrast, Ostrom rejects the inevitability of the tragedy of the commons, advocating polycentric governance that mobilizes global, national, and local actors through iterative policy experimentation [3]. This theory provides key theoretical support for understanding non-compulsory cooperation in global climate governance. The Paris Agreement embodies this tension: its bottom-up design expands participation but relies on voluntary contributions without enforceable sanctions, leaving free-riding unresolved [4]. This means that it is difficult to overcome the collective action dilemma solely through voluntary mechanisms. Ernst et al. propose a differentiated carbon club that combines intergenerational transfers and regional carbon pricing to offset welfare losses and incentivize broad participation [5].

1.3. Research gaps in the incentive mechanism literature

The existing literature has extensively discussed the problem of "hitchhiking" in international climate cooperation and has analyzed the causes and countermeasures to cooperation difficulties from different angles. However, most of these studies are independent and have not provided a systematic answer to a key question: how to integrate the existing decentralized program into a comprehensive incentive mechanism framework that accounts for "sanction effectiveness" and "fair feasibility" simultaneously, thereby promoting international climate cooperation? And this is exactly the core question that this paper tries to answer.

1.4. Analytical framework and contribution

By comparing the incentive structures of the Kyoto Protocol and the Paris Agreement, this paper reveals the root cause of the "incentive incompatibility" in the current climate cooperation mechanism. On this basis, this paper identifies three dilemmas facing the Paris Agreement: hitchhiking incentives remain, there is a lack of effective punishment mechanisms, and short-term benefits continue to overwhelm long-term goals. In response to these problems, this paper proposes three solutions: a climate club and progressive sanctions; a credible commitment and execution

mechanism; and cost-benefit rebalancing, which together form a complete incentive mechanism optimization framework.

2. Institutional evolution: From the Kyoto Protocol to the Paris Agreement

2.1. The Kyoto Protocol: Centralized commitments and compliance failures

The Kyoto Protocol represents the first global attempt to impose legally binding emission reductions. The Kyoto Protocol was reached at the Third Conference of the Parties held in Kyoto, Japan, in December 1997. Adopting a top-down approach, it mandated developed countries to cut greenhouse gas emissions by an average of 5.2% below 1990 levels during the 2008–2012 commitment period. Implementation faltered when the United States withdrew in March 2001, citing economic costs and the absence of developing-country obligations. Canada followed in December 2011, exiting to avoid an estimated CAD 14 billion in compliance costs after emissions rose 17% above 1990 levels. Although Canada has been criticized by international public opinion, it has not been subjected to additional substantive penalties beyond the waiver of its performance obligations. The Kyoto Protocol ultimately failed to curb the continuous upward trend of global emissions.

Table 1. Kyoto Protocol vs. Paris Agreement

Dimension	Kyoto Protocol (1997)	Paris Agreement (2015)
Coverage	37 Annex I countries	194 Parties (>97% emissions)
Targets	Topdown, legally binding	Bottomup, nonbinding NDCs
Period	2008–2012	2020 onward (5year cycles)
Avg. Reduction	–5.2% vs 1990	~–7% by 2030 vs 2019
Carbon Market	ETS, CDM, JI	Article 6 (voluntary)
Transparency	Basic reporting	Enhanced Transparency Framework
Enforcement	Weak compliance body	No formal enforcement
Emission Trend	+15% CO ₂ (2000–2012)	NDCs → 2.5–2.9°C pathway
Temp. Goal	Not defined	≤2°C / pursue 1.5°C
Legacy	Binding precedent	Universal participation

According to Table 1, this comparison highlights a fundamental trade-off: the Kyoto Protocol prioritized legal precision at the cost of limited participation, whereas the Paris Agreement achieved near-universal engagement but faces persistent challenges in ambition, enforcement, and credibility.

2.2. The Paris Agreement: Decentralized pledges and ambition gaps

The 2015 Paris Agreement marked a paradigm shift toward a bottom-up framework of nationally determined contributions (NDCs). Countries set their own emission-reduction targets and action plans, and no longer allocate emission-reduction obligations through international agreements. The agreement also calls for a global inventory every five years to assess collective progress and encourage countries to gradually improve their efforts in subsequent submissions. In addition, the agreement establishes a transparency framework that requires countries to regularly report on emissions data and implementation progress, and supports mechanisms for developed countries to

provide funding and technical support to developing countries. Yet the model suffers from two structural weaknesses: NDCs lack legal binding force, and no enforcement mechanism to penalize non-compliance, leaving free-riding unaddressed. According to UNEP's 2023 Emissions Gap Report, full implementation of current NDCs would still result in 2.5–2.9°C of warming—far above the 1.5°C goal [6].

3. Diagnostic analysis of institutional failures in climate governance

3.1. Institutional contributions and structural limits

3.1.1. The Kyoto Protocol: Institutional legacy and limited impact

The Kyoto Protocol represents a pioneering effort to institutionalize global climate governance. The Kyoto Protocol plays an important role in international climate governance. According to the IPCC AR6 comprehensive report, the agreement has led to reductions in greenhouse gas emissions in some countries and has played a key role in capacity-building for greenhouse gas reporting, accounting, and emission trading markets [1]. It introduces market-oriented tools, such as emissions trading, which provide an institutional basis for the subsequent carbon market system. Although its overall emission reduction is limited, it is a milestone in institutionalizing climate governance.

3.1.2. The Paris Agreement: Universal participation and decentralized governance

The scope of participation in the Paris Agreement is far broader than that of the Kyoto Protocol. According to the Secretariat of the United Nations Framework Convention on Climate Change, the vast majority of countries worldwide have signed and ratified the agreement, covering the vast majority of the world's sources of greenhouse gas emissions. It marks the first inclusion of all major emitters within a single, albeit flexible, mitigation framework. The wide range of participation is closely related to adjustments to the Paris Agreement's mechanism. The agreement abandons the "top-down" mandatory model of the Kyoto Protocol and adopts the "bottom-up" autonomous contribution model instead. Countries submit their own emission-reduction targets based on their national conditions, rather than being set uniformly by international treaties. At the same time, the agreement introduces a global inventory mechanism every five years, requiring countries to gradually raise their level of commitment on this basis. In addition, the establishment of the transparency framework, which requires States to report regularly and be subject to review, partially addresses the lack of mandatory penalties [7].

3.2. Persistent institutional deficits

3.2.1. Enduring free-riding incentives

Neither the "top-down" model of the Kyoto Protocol nor the "bottom-up" model of the Paris Agreement has fundamentally eliminated the hitchhiking incentive.

Under the Kyoto Protocol, the United States announced its withdrawal in 2001, citing that it "damaged the U.S. economy" and that it "had no obligations for developing countries" [1]. This behaviour is a typical case of hitchhiking: the United States avoids domestic emission-reduction costs while continuing to enjoy the benefits of global climate stability. In 2011, Canada also officially withdrew, citing its failure to meet the emission-reduction target without any substantive punishment. These two withdrawals exposed the core flaw of the Kyoto Protocol: when a state finds

that the cost of fulfilling its commitments exceeds the benefits, withdrawal becomes a rational choice.

Under Paris, free-riding persists through systematically modest NDCs rather than outright withdrawal. According to a 2023 United Nations Environment Programme report, even if all countries fulfil their voluntary contribution commitments, the global temperature will still rise by 2.5 to 2.9°C, far above the 1.5 °C temperature control target [6]. There is a clear gap between countries' emission-reduction commitments and the reductions required to achieve the temperature-control target. This shows that countries generally choose lower standards when setting targets to reduce their own emissions, while hoping that other countries will take on more responsibility. In other words, the voluntary commitment mechanism has failed to effectively change each country's income structure.

3.2.2. Absence of credible enforcement mechanisms

Both the Kyoto Protocol and the Paris Agreement have failed to establish an effective mechanism for punishment, which is a common institutional flaw.

Although the Kyoto Protocol has established a compliance mechanism, its punishment authority is very limited. Although the mechanism has a compliance committee, it lacks the power to impose sanctions, including trade sanctions or economic fines, on the defaulting country. Its accountability is mainly limited to technical measures and procedural supervision, and its substantive binding force is extremely weak. This means that when a country fails to meet its emission-reduction target, the consequences are not serious. Canada finally officially withdrew from the agreement in 2011 because it was unable to fulfil its emission-reduction target for the first commitment period and faced significant performance costs, but did not incur any economic penalties. After the United States announced its refusal to ratify in 2001, it was also not subject to any substantive sanctions from the international community.

Paris relies exclusively on self-reporting and periodic stocktaking, with no punitive provisions whatsoever. If a country fails to achieve its goal of voluntary contributions, the most serious consequence is being publicly discussed or criticized at international conferences. This institutional design means that the cost of betrayal of cooperation is almost zero.

3.2.3. Temporal mismatch between costs and benefits

A deeper structural barrier is temporal mismatch: immediate abatement costs contrast with deferred climate benefits.

Under the Kyoto Protocol, the typical manifestation of this problem is that some industrialized countries struggle to meet their emission-reduction obligations. For example, Canada became the first country to withdraw from the Kyoto Protocol in 2011. Despite widespread protests from the international community, domestic economic priorities still outweigh international climate commitments. Under the Paris Agreement, short-termism is reflected in the general tendency to set lower standards for countries when setting independent contribution goals. The deep reason is that climate mitigation measures require large-scale investment in the short term. The benefits of stabilizing the climate can only be realized in the long term - this time mismatch between costs and benefits makes it difficult for the government to convince voters of the rationality of high early expenditure, especially given the short election cycle [7]. When short-term economic growth conflicts with long-term climate goals, political rationality tends to favour the former.

4. Designing incentive-compatible reforms for global climate governance

4.1. A dual-track hybrid governance architecture

Building on Manolas's critique of institutional asymmetry, this section proposes a dual-track hybrid governance framework that reconciles legitimacy with enforceability. According to Manolas, the Kyoto Protocol has rules but lacks universal participation, and the Paris Agreement has participation but lacks hard binding [8]. This article proposes to build a "double-layer hybrid governance" framework to combine the advantages of both. The universal participation framework of the Paris Agreement will be retained at the top level, and the country's independent contribution will continue to be the basis for its commitment to reduce emissions and to recognize differences in countries' stages of development. Introduce the core of quasi-mandatory rules at the bottom, legalize core clauses such as MRV, carbon market rules, and a minimum carbon price range, and unify performance procedures and accounting standards. At the same time, the NDC upgrade channel is designed to allow the voluntary commitment to be converted into a binding sub-agreement after review, thereby reducing the likelihood of withdrawal. Crucially, this upgrade mechanism introduces a soft path from voluntarism to legal bindingness, reducing strategic under-reporting at the design stage.

4.2. Climate clubs and graduated enforcement

To change the game with the income structure of hitchhiking, this article proposes establishing a climate club, initiated by major emitting economies, to set a unified minimum price range and implement border carbon adjustment for non-members. At the same time, a progressive sanctions mechanism is introduced: the first stage is international review and public rating; the second stage restricts access to green financing and technology transfer; and the third stage is the implementation of trade measures. To prevent club division, a transitional support mechanism should be established to link sanctions to the degree of effort rather than just the results. This effort-based linkage mitigates North–South equity concerns and lowers entry barriers for emerging economies.

4.3. Credible commitment and enforcement mechanisms

To address the lack of a punishment mechanism, this paper proposes establishing a domestic law docking mechanism to encourage countries to incorporate NDC targets into domestic legislation and to give domestic courts the power to review climate performance. Such domestic entrenchment transforms international commitments into justiciable obligations, raising the political cost of non-compliance. At the same time, an automatic trigger mechanism should be designed to launch a supplementary emission-reduction plan or international review procedures if a country fails to meet the stage target. In addition, strengthen the executive authority of the existing compliance committee, introduce independent third-party verification agencies, and directly link the verification results to financing and market access. This linkage aligns climate performance with material economic benefits, creating endogenous enforcement incentives.

4.4. Cost–benefit rebalancing and intertemporal incentives

According to Cash and Swatuk, there is a serious performance gap in global climate funds at present. Most developed countries have not fulfilled their capital commitments, and climate funds are mostly provided as loans rather than grants, thereby increasing the debt burden of developing

countries [9]. This structural imbalance underscores the need for innovative financial architectures that decouple climate action from short-term fiscal vulnerability. According to Emam, climate governance is generally short-sighted. Countries pay more attention to short-term economic interests and ignore long-term climate responsibilities and sustainable development [10]. Based on this, this paper proposes establishing an intergenerational transfer payment mechanism to provide long-term, stable financial support from high-income countries to low-income countries with fragile climates, thereby alleviating the welfare losses caused by energy transformation. At the same time, climate spending is regarded as a long-term investment in infrastructure and industrial upgrading, relying on multilateral development banks and green bonds to relieve short-term fiscal pressure, and binding climate policies with real benefits such as employment and energy security to hedge against the shortfall of insufficient long-term returns with immediate benefits. By aligning climate policy with visible domestic dividends, governments can sustain public support across electoral cycles.

5. Conclusion

5.1. Main findings

Starting from the dilemma theory of collective action, this article compares the two governance models of the Kyoto Protocol and the Paris Agreement and identifies three core problems: the long-term persistence of "hitchhiking", the lack of effective punishment, and the short-term interests that overwhelm long-term goals. Based on this, the article offers suggestions from three perspectives. First, establish climate clubs and progressive sanctions. The minimum carbon price is set by the major emitting economies, border carbon adjustment is implemented for non-members, and graded sanctions, ranging from public ratings to trade measures, are adopted. Second, build a credible commitment and execution mechanism. Promote countries to include emission reduction targets in domestic legislation, introduce automatic triggering mechanisms, and strengthen the authority of the compliance committee and third-party verification. Third, promote rebalancing of the cost-benefit ratio. Through intergenerational transfer payments, green bonds, and multilateral development banks, climate policies are linked to real benefits such as employment and energy security.

5.2. Theoretical and practical contributions

The research significance of this article lies at two levels. At the social level, the incentive framework helps alleviate the dilemma of "the willingness to participate and the hitchhiking incentive coexist" in climate cooperation, reduces the risk of cooperation failure due to a lack of trust, and provides a workable institutional design for international climate negotiations. At the commercial level, this article provides clear policy expectations for enterprises, reduces the uncertainty in green technology research and development and carbon market transactions, and helps to improve investors' confidence in low-carbon projects such as renewable energy, energy storage and carbon capture, so as to promote the large-scale development of related industries. In general, this article provides a theoretical framework for policymakers to design climate cooperation mechanisms and an analytical basis for enterprises to predict future trends in carbon policies.

5.3. Limitations and avenues for future research

This research is primarily based on comparisons of second-hand information and systems and lacks first-hand data support. While this approach ensures breadth of institutional comparison, it

inevitably constrains empirical granularity and causal identification. In the future, the reliability and applicability of the conclusions can be further enhanced through methods such as interviews, game simulations, or case assessments (such as the EU's carbon border adjustment mechanism). Such extensions would enhance the external validity of the incentive-compatible framework and inform the calibration of real-world climate institutions.

References

- [1] Bongaarts, J. (2024). IPCC, 2023: Climate Change 2023: Synthesis Report. *Population and Development Review*, 50(2), 577–580.
- [2] Nordhaus, W. (2015). Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, 105(4), 1339–1370.
- [3] Ostrom, E. (2017). Polycentric systems for coping with collective action and global environmental change. In *Global Justice* (pp. 423–430). Routledge.
- [4] Doelle, M. (2026). The paris agreement: Historic breakthrough or high stakes experiment? *Climate Law*, 6(1–2), 1–20.
- [5] Ernst, A., Hinterlang, N., Mahle, A., & Stähler, N. (2023). Carbon pricing, border adjustment and climate clubs: Options for international cooperation. *Journal of International Economics*, 144, 103772.
- [6] Emissions Gap Report 2023: Broken Record – Temperatures hit new highs, yet world fails to cut emissions (again). (2023). United Nations Environment Programme.
- [7] Falkner, R. (2026). The Paris Agreement and the new logic of international climate politics. *International Affairs*, 92(5), 1107–1125.
- [8] Manolas, E. (2016). The Paris climate change agreement. *International Journal of Environmental Studies*, 73(2), 167–169.
- [9] Cash, C., & Swatuk, L. A. (2022). Climate finance: Lessons from development finance. In *International Political Economy Series* (pp. 1–19). Springer International Publishing.
- [10] Emam, H. A. (2025). Climate change and monetary policy. In *Encyclopedia of Monetary Policy, Financial Markets and Banking* (pp. 26–31). Elsevier.